DOCUMENT RESUME

ED 207 555 IR 009 667

AUTHOR	Siegel, Martin A.; DiBello, Louis V.
TITLE	Optimization of Computerized Drills: An Instructional Approach.
INSTITUTION	Illinois Univ., Urbana. Computer-Based Education Research Lab.
SPONS AGENCY	Army Research Inst. for the Behavioral and Social Sciences, Alexandria, Va.
REPORT NO	CERL-R-E-16
PUB DATE	Apr 80
CONTRACT	MDA903-79-C-0551
NOTE	39p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 1980).
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Citizenship Education; College Students; *Computer Assisted Instruction; Decoding (Reading); *Discrimination Learning; *Drills (Practice); *Feedback; Instructional Design; *Microcomputers; *Models; Reading Comprehension
IDENTIFIERS	PLATO

ABSTRACT

This report describes a computer drill called the Corrective Feedback Paradigm (CFP) and summarizes a research study which examines the effects of the CFP on posttest and time to mastery achievement. The CFP, which embodies the use of such principles as mastery learning, increasing ratio review, and discrimination training, is capable of being implemented on any computer-assisted instruction system, including micro-based systems. In this analysis of the program, 120 college seniors were required to learn 20 English-Japanese transliterated word pairs. Treatment groups were then examined according to the type of corrective feedback given, and the type of training and review procedures used. Results indicated that increasing numbers of ratio review techniques should be employed in computer drills, and that discrimination training should be provided after error feedback. Examples of CFP drills on reading comprehension and constitutional history are included, as well as recommendations for the future design of optimized drill paradigms. (MER)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED207555

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Optimization of Computerized Drills: An Instructional Approach

Martin A. Siegel

Computer-based Education Research Laboratory
and the Department of Educational Psychology
University of Illinois at Urbana-Champaign

Louis V. DiBello

Computer-based Education Research Laboratory
University of Illinois at Urbana-Champaign

Paper presented at the Annual Meeting of the
American Educational Research Association

Boston, Massachusetts

April, 1980

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. Bitzer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

IR009667

Copyright © April 1980
by the Board of Trustees, University of Illinois

All rights reserved. No part of this book may be reproduced in any form or by any means without permission in writing from the authors

This work was partially supported by the Army Research Institute for the Behavioral and Social Sciences under contract MDA903-79-C-0551.

Optimization of Computerized Drills: An Instructional Approach

Drills have been employed by educators for hundreds of years. Drill and practice routines in one form or another are used in nearly every aspect of education. Students practicing their arithmetic math facts, reading words printed on flash cards, rehearsing the spelling of a list of words, or memorizing a foreign language vocabulary are common examples of drills used in schools. Typically these drills are handled in the following way: a stimulus is written on one side of a card (e.g., "6 + 3"); the appropriate response is written on the other side of the card (e.g., "9"). The instructor starts the drill by shuffling the stack of cards and showing the top card. The student looks at the stimulus and says the response; the card is turned over to provide feedback. If the student is correct, the card is removed from the stack; if the student is incorrect, the card is placed at the end of the stack. When all cards are removed from the stack, the drill is over.

In many ways the computer is an ideal medium for presenting drill and practice instruction. The computer is infinitely patient -- i.e., it can present items until a mastery criterion is met. The computer can provide consistent and immediate feedback for correct and incorrect responses; it can branch to remedial or advanced material; and it can adjust the pace according to the needs of the individual student. The computer can store the student's performance data and keep track of the student's place in a complicated curriculum sequence. Moreover, the computer can modify the items of a drill according to prespecified parameters, so that the student cannot use false prompts as a learning crutch.

An important instructional question is how to design a more effective and efficient drill. How can drills be constructed so as to minimize time to mastery and maximize retention? This question is especially critical in

the design of computerized drill programs, for the computer can only behave according to the scheme explicitly provided by the instructional designer.

A significant theoretical step in addressing this question was taken by Atkinson (1972a, 1972b) in applying the techniques of optimal control theory (Boltyanskii, 1971; Kalman, Falb, & Arbib, 1969) and simple stochastic learning models (Atkinson & Crothers, 1964; Calfee & Atkinson, 1965) to the optimization of item sequencing algorithms for paired-associate list drills. The optimized drill routine which Atkinson developed from this learning model (the so-called optimized-unequal parameter case) performed significantly better than the other drills. The average score on a delayed posttest administered one week after treatment was 79% for the optimized drill group, whereas it was 58% for a drill routine in which the student selected the next item in the drill and 38% for a random order drill (Atkinson, 1972b). These impressive results indicate that mathematical optimization techniques can produce effective drill algorithms -- even when the underlying mathematical model is very simple.

An entirely different approach was taken by Siegel and Misselt (1974) with the development of the "Corrective Feedback Paradigm" (CFP). This drill paradigm embodied a number of instructional principles such as mastery learning, increasing ratio review, and discrimination training. These features were based on wide experience in teaching and on the development and implementation of instructional materials. Corroboration for the CFP's features can be found in the psychological research literature on feedback (Kulhavy, 1977) and in paired-associate item spacing studies (e.g., Greeno, 1964; Hintzman, 1974; Keppel, 1964; Peterson, 1966; Peterson, Wampler, Kirkpatrick, & Saltzman, 1963; Young, 1971). The drill paradigm has been in use for several years on the PLATO Computer-based Education System (Smith & Sherwood, 1976), although it is not dependent on any particular hardware or sorftware configuration. The CFP is capable of being implemented on any computer-assisted instruction system -- including micro-based systems.

The purpose of this paper is to describe the CFP, to summarize a research study which examined the effects of the CFP on posttest achievement and time to mastery, to present three examples of the CFP in instructional settings, and to discuss the future design of optimized drill paradigms.

The Corrective Feedback Paradigm

In order to describe the paradigm's main features, an example of a drill will be presented. Some instructional questions will be asked about the drill, and the CFP's features will be shown to address these questions.

Suppose an instructor wanted to design a computerized drill to teach the names for the 103 symbols in the periodic table of the elements (H is hydrogen, He is helium, Li is lithium, etc.). Of course, not all of the 103 elements would be presented during a single drill session, since that would be too difficult for most students. Rather, the 103 items in the drill (each item being a symbol-name pair) would be subdivided into a sequence of several smaller drills. For purposes of simplification, only the instructional features of one of these smaller drills will be described.

In this example, a drill item is presented by displaying a stimulus (e.g., the symbol "Si") on the computer screen. A pointer or arrow (>) indicates that the student should type a response:

Si

>

After the student types a response and indicates (by pressing a special key on the computer keyset) that the response should be judged, feedback is provided (e.g., "ok" or "no" or some other form of feedback):

Si

silicon ok

This is the basic format of a computerized drill. A stimulus is presented; the student types a response; the response is judged as correct or incorrect; and appropriate feedback is provided. However, a number of questions must still be answered before such a drill can be programmed. What kind of feedback does the computer provide the student if a mistake is made? Are all mistakes treated in the same way? What criteria is used to determine when an item is "mastered" and should be retired (or removed) from the drill? Will a missed item be reviewed; and if so, according to what schedule? Various features of the CFP were designed to answer these instructional questions. The following briefly describes these features.

Type of Corrective Feedback

The type of corrective feedback provided by the CFP is a function of the type of mistake made. If the student types a response which is not an answer to another item in the drill list -- an out-of-list error (OLE) -- the correct answer is provided. For example, if the student types "metal" for the stimulus "Si," and if "metal" is not an answer to another drill item (i.e., is not a name for one of the 103 chemical elements in the periodic table), the paradigm provides the correct answer:

OLE:	OLE Feedback:
Si	Si
> metal	metal no
	Si is silicon

If the student types a response which is an answer to another item in the drill list -- a discrimination error (DE) -- discrimination feedback is provided. For example, if the student types "silver" (the answer for the

stimulus "Ag") instead of "silicon" (the answer for the stimulus "Si"), DE feedback is provided:

DE:	DE Feedback:
Si	Si
> silver	silver is Ag
	Si is <u>silicon</u>

Discrimination feedback tells what stimulus the student responded to as well as the answer to the stimulus that was actually presented.[1]

<u>Discrimination Training</u>

If the student makes a discrimination error (an intra-list confusion), the CFP provides discrimination feedback (e.g., "silver is Ag; Si is <u>silicon</u>") as described above. In addition, the student is given discrimination <u>training</u> -- i.e., the student is simultaneously presented the <u>missed</u> stimulus and the <u>confused</u> stimulus. For example, the student is asked to type the name for the symbols "Si" <u>and</u> "Ag." Both items are presented at the same time, so that the student can study their similarities and differences:

Discrimination Training:
(Both stimuli displayed at once)

Si	Si	Si
>	silicon ok	silicon ok
Ag	Ag	Ag
	>	silver ok

Discrimination training is completed when both items (presented in random order) are answered correctly. Of course, if a mistake is made during

discrimination training, OLE or DE feedback is provided. During discrimination training, new DEs (e.g., confusing "Ag" for "argon") are treated as OLEs because it would be confusing for the student to receive training on more than one DE at a time.

Retirement Criterion

If an item (a stimulus-response pair such as Si-silicon) is correctly answered, it is placed at the end of the list or stack of items in the drill. Then the next item in the queue is presented to the student. An item is retired (dropped) from the list of items when it has been answered correctly k consecutive times, where k is specified by the designer of the drill. For example, if k is 4, each item must be answered correctly four times without an intervening mistake on that item.

Increasing Ratio Review

Traditional computerized drills emulate the procedures used in non-computerized drills; namely, an incorrectly answered item is placed at the end of the stack of items to be presented, and a correctly answered item is removed from the stack. However, there are two problems with this procedure for reviewing a missed item: 1) if there are many items left in the stack of items to be presented, the correct response for the missed item probably will be forgotten by the time the missed item reappears for review; and 2) if there are few items left in the stack of items to be presented, the correct response for the missed item probably will be remembered by the time the missed item reappears for review, but the student will not have been "shaped" into remembering the correct response over a longer time period. The CFP addresses both of these problems by using an increasing ratio review schedule.

After the student makes a mistake and receives either form of corrective feedback (and possibly discrimination training), the missed item reappears

according to a spaced review schedule determined by the designer's specification of review parameters. The instructor may specify up to three review positions by assigning a value to the parameters review position #1, review position #2, and review position #3 (where review position #1 < review position #2 < review position #3). For example, if review position #1 is 3, review position #2 is 6, and review position #3 is 10, then the first review of the missed item would occur three item presentations after the error was made; if that review were answered correctly, the second review would be given six item presentations after the first review; and if that review were answered correctly, the third review would be given ten item presentations after the second review. At any stage during this review schedule, an incorrect response on a reviewed item would reinitiate the review sequence (beginning with review position #1) for that item.

An example will demonstrate this increasing ratio review procedure. Imagine a drill with 14 items, a through n (e.g., 14 randomly selected symbol-element name pairs). Suppose the retirement criterion is 2; that is, an item is "dropped" from the list of items when the item is answered correctly two times in a row without an intervening mistake.

A hypothetical student's performance is shown in Figure 1. Each row of this figure shows how the queue (or stack) of items was arranged at the time of each item presentation. The first column of each row (queue position #1) shows the specific item that was presented to the student. By reading down the first (shaded) column, one can determine the exact sequence of items that was presented to our hypothetical student. Thus, as seen in Figure 1, the student was presented item a, then item b, then item c, then item b again, then item d, and so on.

Insert Figure 1 about here

ITEM QUEUES FOR PRESENTATIONS 1 - 22

PRESENTATION #	QUEUE POSITION 1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	a	b	c	d	e	f	g	h	i	j	k	l	m	n
2	B	c	d	e	f	g	h	i	j	k	l	m	n	a
3	c	(b)	d	e	f	g	h	i	j	k	l	m	n	a
4	(b)	d	e	f	g	h	i	j	k	l	m	n	a	c
5	d	e	f	(b)	g	h	i	j	k	l	m	n	a	c
6	e	f	(b)	g	h	i	j	k	l	m	n	a	c	d
7	f	(b)	g	h	i	j	k	i	m	n	a	c	d	e
8	(b)	g	h	i	j	k	l	m	n	a	c	d	e	f
9	g	h	i	j	k	(b)	l	m	n	a	c	d	e	f
10	h	i	j	k	(b)	l	m	n	a	c	d	e	f	g
11	I	j	k	(b)	l	m	n	a	c	d	e	f	g	h
12	j	k	(b)	l	m	n	a	c	d	e	f	g	h	i
13	k	(b)	l	m	n	a	c	d	e	f	g	h	i	j
14	(b)	l	m	n	a	c	d	e	f	g	h	i	j	k
15	l	m	n	a	c	d	e	f	g	h	i	j	k	b
16	M	n	a	c	d	e	f	g	h	i	j	k	b	l
17	n	(m)	a	c	d	e	f	g	h	i	j	k	b	l
18	(m)	a	c	d	e	f	g	h	i	j	k	b	l	n
19	a	c	d	(m)	e	f	g	h	i	j	k	b	l	n
20	c	d	(m)	e	f	g	h	i	j	k	b	l	n	
21	d	(m)	e	f	g	h	i	j	k	b	l	n		
22	(M)	e	f	g	h	i	j	k	b	l	n			

ITEM QUEUES FOR PRESENTATIONS 23 - 43

PRESENTATION #	QUEUE POSITION 1	2	3	4	5	6	7	8	9	10	11	12	13	14
23	e	(m)	f	g	h	i	j	k	b	l	n			
24	(m)	f	g	h	i	j	k	b	l	n				
25	f	g	h	(m)	i	j	k	b	l	n				
26	g	h	(m)	i	j	k	b	l	n					
27	h	(m)	i	j	k	b	l	n						
28	(m)	i	j	k	b	l	n							
29	I	j	k	b	l	(m)	n							
30	j	(i)	k	b	l	(m)	n							
31	(i)	k	b	l	(m)	n								
32	k	b	l	(i)	(m)	n								
33	b	l	(i)	(m)	n									
34	l	(i)	(m)	n										
35	(i)	(m)	n											
36	(m)	n	b	a	c	(i)								
37	n	b	a	c	(i)	m								
38	b	a	c	(i)	m									
39	a	c	(i)	m										
40	c	(i)	m											
41	(i)	m												
42	m	i												
43	i													

Figure 1. Drill History of a Hypothetical Student

In this figure, a lower case letter in the first queue position indicates that the item was answered correctly, and a capital letter in the first queue position indicates that the item was missed. If every item presentation were answered correctly, the items always would be placed at the end of the queue, and the entire queue would be cycled through twice (since the retirement criterion is 2) with no change in the order of the items. But an incorrect response on an item initiates the increasing ratio review schedule, and this does affect the order of the items.

Figure 1 will illustrate how the item queue changed dynamically according to our hypothetical student's responses. For this example, let us assume the instructor preselected review position #1 = 2, review position #2 = 4, and review position #3 = 6. If an item is answered correctly (e.g., item a at presentation #1), it is "moved" to the end of the queue. But, if an item is missed (e.g., item b at presentation #2 was missed and was written, therefore, B), it is moved to the second queue position for "2-later review"; if the item is answered correctly when it reappears for 2-later review it is moved to the fourth queue position for "4-later review"; if the item is answered correctly when it reappears for 4-later review it is moved to the sixth queue position for "6-later review." At any time during this review schedule, if an item is missed again, it is moved to the second queue position for 2-later review; i.e., a mistake during 2-, 4-, or 6-later review will "restart" the review schedule for that item to 2-later review.

This "rearranging of the queue" is analogous to inserting an incorrectly answered flash card in the middle of a stack of flash cards. For 2-, 4-, 6-later review, the procedure would work as follows. If the item appearing at the top of the stack is missed, it is inserted after the next card in the stack (or two cards later) instead of being placed at the end of the stack as in a traditional drill. When the missed flash card reappears for review, it either is answered correctly or incorrectly; if it is answered correctly, the card is inserted after the third card in the stack (or four

cards later); but, if it is answered incorrectly, the review schedule is "restarted" so the card is inserted after the next card in the stack (or two cards later). Similarly, this procedure is followed when the card reappears for 4-later review; if it is answered correctly, the card is inserted after the fifth card in the stack (or six cards later); if it is answered incorrectly, the card is inserted after the next card in the stack (or two cards later). (Note that each row of Figure 1 would represent the arrangement of the stack of 14 flash cards at any given point in time.) Although one could conceive of this procedure being implemented with flash cards, in practice it would be unbearably difficult since the student can make a mistake on more than one item and the review schedules therefore would interact. Fortunately, computers easily can handle this type of bookkeeping task!

Returning to Figure 1 (segments of the figure are repeated in the text below), notice what happened. Item a was answered correctly by our hypothetical student on presentation #1. If we examine the queue for presentation #2, we note that 1) item b was the next item presented; 2) item b was answered incorrectly (thus written B); and 3) item a had been placed at the end of the queue.

PRESENTATION	QUEUE POSITION													
#	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	a	b	c	d	e	f	g	h	i	j	k	l	m	n
2	B	c	d	e	f	g	h	i	j	k	l	m	n	a

Then item c appeared as presentation #3 and it was answered correctly. Note that item b had been placed in the second queue position for its 2-later review. (In Figure 1 and below, reviewed items are circled so that their history can be followed more easily.) Then item b reappeared as presentation #4 (2-later review); it was answered correctly.

PRESENTATION	QUEUE POSITION													
#	1	2	3	4	5	6	7	8	9	10	11	12	13	14
3	c	(b)	d	e	f	g	h	i	j	k	l	m	n	a
4	(b)	d	e	f	g	h	i	j	k	l	m	n	a	c

Then items d, e, and f were answered correctly and placed at the end of the queue. Item b reappeared (4-later review) as presentation #8 and again six items later (6-later review) as presentation #14.

PRESENTATION #	QUEUE POSITION 1	2	3	4	5	6	7	8	9	10	11	12	13	14
5	d	e	f	(b)	g	h	i	j	k	l	m	n	a	c
6	e	f	(b)	g	h	i	j	k	l	m	n	a	c	d
7	f	(b)	g	h	i	j	k	l	m	n	a	c	d	e
8	(b)	g	h	i	j	k	l	m	n	a	c	d	e	f
9	g	h	i	j	k	(b)	l	m	n	a	c	d	e	f
10	h	i	j	k	(b)	l	m	n	a	c	d	e	f	g
11	i	j	k	(b)	l	m	n	a	c	d	e	f	g	h
12	j	k	(b)	l	m	n	a	c	d	e	f	g	h	i
13	k	(b)	l	m	n	a	c	d	e	f	g	h	i	j
14	(b)	l	m	n	a	c	d	e	f	g	h	i	j	k

Since item b was answered correctly during presentation #14, it was moved to the end of the queue and item l (presentation #15) appeared. Presentation #16, (M), was missed, and it was reviewed (2-later review) during presentation #18.

PRESENTATION #	QUEUE POSITION 1	2	3	4	5	6	7	8	9	10	11	12	13	14
15	l	m	n	a	c	d	e	f	g	h	i	j	k	b
16	M	n	a	c	d	e	f	g	h	i	j	k	b	l
17	n	(m)	a	c	d	e	f	g	h	i	j	k	b	l
18	(m)	a	c	d	e	f	g	h	i	j	k	b	l	n

Item a reappeared at presentation #19. It was answered correctly for the second time in a row without an intervening mistake, so item a was dropped from the queue. Similarly, items c and d (presentations #20 and #21, respectively) were answered correctly, so they were dropped from the queue as well. For the 4-later review, item m was answered incorrectly again (presentation #22). Consequently, the 2-, 4-, 6-later review process started over for item m. However, before this review process for item m was completed, item i was answered incorrectly (I) during presentation #29.

Notice that the 2- and 4-later reviews for item i took precedence over the 6-later review for item m (because the error on item I was more recent). So the 6-later review for item m was "pushed back" in the queue. Thus the second and third review position parameters (in this case, 4- and 6-later review positions) actually specify minimum spacings for the last two reviews of missed items.

PRESENTATION #	QUEUE POSITION 1	2	3	4	5	6	7	8	9	10	11	12	13	14
19	a	c	d	(m)	e	f	g	h	i	j	k	b	l	n
20	c	d	(m)	e	f	g	h	i	j	k	l	l	n	
21	d	(m)	e	f	g	h	i	j	k	b	l	n		
22	(M)	e	f	g	h	i	j	k	b	l	n			
23	e	(m)	f	g	h	i	j	k	b	l	n			
24	(m)	f	g	h	i	j	k	b	l	n				
25	f	g	h	(m)	i	j	k	b	l	n				
26	g	h	(m)	i	j	k	b	l	n					
27	h	(m)	i	j	k	b	l	n						
28	(m)	i	j	k	b	l	n							
29	I	j	k	b	l	(m)	n							
30	j	(i)	k	b	l	(m)	n							
31	(i)	k	b	l	(m)	n								
32	k	b	l	(i)	(m)	n								
33	b	l	(i)	(m)	n									
34	l	(i)	(m)	n										

At presentation #33, item b was answered correctly for the second time in a row -- at presentations #14 and #33 -- so it was dropped from the queue. (As an operational rule, correct responses during the first two review positions do not count toward the retirement criteria of two consecutive correct responses.)

Note that at presentation #35, (i), there were three items in the queue: i, m, and n. Since item i was answered correctly during presentation #35 (4-later review), it had to be placed six items later for 6-later review. But there were only three items in the queue! At this point, the paradigm "resurrected" three previously retired items so that review spacing was guaranteed. Thus, at presentation #36, (m), there were six items in the queue: m, n, b, a, c, and i.

PRESENTATION #	QUEUE POSITION 1	2	3	4	5	6	7	8	9	10	11	12	13	14
35	(i)	(m)	n											
36	(m)	n	b	a	c	(i)								
37	n	b	a	c	(i)	m								
38	b	a	c	(i)	m									
39	a	c	(i)	m										
40	c	(i)	m											
41	(i)	m												
42	m	i												
43	i													

All of these features of the CFP -- type of feedback, discrimination training, retirement criterion, and increasing ratio review -- are adjustable for each individual application. For example, the paradigm is presently being used on the PLATO computer-based system for presenting drills on medical terminology, spelling, arithmetic facts, English language vocabulary, U.S. and Illinois State Constitutions, reading comprehension skills, foreign language vocabulary, English as a second language, identification of bones, and others.

These implementations are quite different from one another. Screen displays were designed for each separate application; thus different implementations do not give the outward impression that they are variations on the same theme. Parameters like the retirement criterion and the review schedules for items that are missed were set for each separate application. The features of discrimination feedback and training are options that were selected according to the needs of the particular application. For example, the foreign language and medical terminology drills employ the special training feature for discrimination errors and they also utilize another special option for handling synonyms.

Although most of the developmental work on the CFP has been carried out in a naturalistic environment, an experimental study has been completed (Siegel & Misselt, 1974). The main findings were the verification of the

effectiveness of the features discussed above such as increasing ratio review, and discrimination feedback and training. Because this research was described only in an unpublished report, it will be summarized here.

A Summary of the Siegel and Misselt Study

The purpose of the CFP study was to examine, within the context of a computer-assisted instruction drill program, the effects of several types of corrective feedback and review procedures on the student's (a) posttest achievement and (b) time to mastery.

The task consisted of learning 20 English - Japanese (transliterated) word pairs -- an unfamiliar task to the subjects. The words were selected so as to maximize semantic, phonologic, and orthographic similarities (e.g.: "like-suki," "dislike-kirai," "nice-kirei," "hateful-nikui," "beef-niku").

Design

There were six treatment groups which varied along three dimensions (type of corrective feedback, type of training, and type of review):

	Increasing Ratio Review	Immediate Review Only	No Review
Discrimination Training	1 Adaptive Feedback	2 Adaptive Feedback	
No Training	3 Adaptive Feedback	4 Adaptive Feedback	5 Adaptive Feedback
No Training	6 Fixed Feedback		

For groups 1 to 5, the type of corrective feedback following an error was contingent on the type of error ("adaptive"). If the student made an OLE, the answer was provided. For example, if the student typed "omama" for "mother" and "omama" was not a response to another stimulus in the list, OLE feedback was "mother is okaasan." If the student made a discrimination (or intralist) error, DE feedback was provided. For example, if the student typed "obasan" for "mother," discrimination error feedback was "obasan is aunt; mother is okaasan."

For group 6, corrective feedback was the same ("fixed") for OLEs and DEs. In each case the correct answer was provided (eg: "mother is okaasan").

For groups 1 and 2, discrimination training always followed discrimination error feedback. That is, if two English words were confused they were displayed simultaneously (in random order) and the student was required to type the Japanese equivalent for each stimulus. For groups 3 to 6 discrimination training never followed the corrective feedback of discrimination errors.

For groups 1, 3, and 6, missed items were reviewed on an increasing ratio schedule (1-, 3-, 5-later review). For groups 2 and 4, missed items were reviewed immediately (1-later review). For group 5, no review was provided; after receiving corrective feedback the student proceeded to the next item.

For all groups, each item was "retired" from the queue when it was answered correctly on two consecutive presentations (excluding review positions #1 and #2) of that item.

Method

The Ss were 102 college seniors enrolled in an educational psychology course. Participation in the experiment fulfilled a course requirement.

Seventeen Ss were randomly assigned to each of the six experimental groups. None of the Ss had prior coursework in Japanese, nor were any familiar with the language.

The entire experiment was controlled by the University of Illinois PLATO IV computer-based educational system. Under every condition the S first received a short on-line orientation to the keyset and general instructions on interacting with PLATO. Then every S was administered a typing ability test. Thirty unfamiliar foreign language (non-Japanese) vocabulary words were individually displayed and the S copied (typed) the word and pressed the "NEXT" key. The S then was presented with a randomly ordered list of the 20 English-Japanese word pairs. The list was displayed for two minutes and the S was instructed to learn as many word pairs as possible during the allotted time. Following the display, each English word was presented individually and the S was instructed to type the Japanese equivalent or press the NEXT key if unable to provide an answer. After this pretest, the S was informed that an instructional drill session was to begin which would enable the S to master the Japanese words.

The instructional drill involved a series of self-paced discrete presentations. Each presentation began by displaying an English word on the screen. The S was required to type the Japanese (transliterated) equivalent or press the "ANS" key (for the answer) if unable to provide a translation. If the S correctly translated the English word, the drill indicated "ok" adjacent to the S's response; the next item in the queue then was presented. If the S asked for the answer or incorrectly responded, corrective feedback, discrimination training, and/or review procedures were initiated according to the S's treatment group. All Ss were instructed to study each feedback message. Depending on the group, the S received DE or OLE feedback contingent on the type of mistake (groups 1 to 5) or OLE feedback without regard to the type of mistake (group 6); discrimination training (groups 1 to 2) or no discrimination training (groups 3 to 6); and increasing ratio review (groups 1, 3 and 6), immediate

review only (groups 2 and 4), or no review (group 5). Each item was "retired" from the queue when it was successfully answered on two consecutive presentations (excluding review positions #1 and #2) of that item. When all of the items were retired (mastery) the drill session ended. The S then answered some attitude questions, took an on-line English vocabulary test (48 items), and then took a posttest on the English-Japanese pairs. Unlike the pretest, an answer was required for each item on the posttest.

Results

Table 1 shows the observed group means, standard deviations, and ranges for three dependent measures: (a) number of DEs on the posttest; (b) number of OLEs on the posttest; and (c) time to criterion (instructional time). (Mean scores on three covariates -- typing test score, vocabulary score, and pretest score -- are not displayed.)

Insert Table 1 about here

For purposes of the analyses, the DE and OLE scores were transformed using the reciprocal transformation $X' = 1/(X+1)$, since the square of treatment means and standard deviations were proportional.

The analysis of variance for the three dependent variables showed: (a) significant differences among the six treatment groups for the mean number of DEs on the posttest, $F\ (5,96) = 4.05$, $p < .01$; (b) significant differences among the six treatment groups for the mean number of OLEs on the posttest, $F\ (5,96) = 9.72$, $p < .001$; and (c) no significant differences

Table 1

Group Means, Standard Deviations, and Ranges for Three Dependent Measures (Untransformed)

Group	Features		DEs	OLEs	Instructional Time (minutes)
1	Adaptive Feedback	Mean:	0.294	1.059	35.020
	Discrimination Training	S.D.:	0.470	1.197	17.736
	Increasing Ratio Review	Range:	0 - 1	0 - 4	18.020 - 83.080
2	Adaptive Feedback	Mean:	2.412	2.824	39.801
	Discrimination Training	S.D.:	2.551	1.185	15.822
	Immediate Review Only	Range:	0 - 8	0 - 4	23.850 - 85.680
3	Adaptive Feedback	Mean:	1.294	1.059	42.888
	No Discrimination Training	S.D.:	1.312	1.345	17.237
	Increasing Ratio Review	Range:	0 - 4	0 - 5	15.460 - 79.540
4	Adaptive Feedback	Mean:	2.353	3.000	34.651
	No Discrimination Training	S.D.:	2.149	1.696	15.229
	Immediate Review Only	Range:	0 - 8	1 - 8	12.970 - 77.070
5	Adaptive Feedback	Mean:	2.235	3.353	32.102
	No Discrimination Training	S.D.:	2.107	2.060	13.292
	No Review	Range:	0 - 8	0 - 7	10.770 - 61.840
6	Fixed Feedback	Mean:	1.470	0.882	33.272
	No Discrimination Training	S.D.:	1.700	1.054	11.745
	Increasing Ratio Review	Range:	0 - 5	0 - 4	17.790 - 66.160

among the six treatment groups for the mean number of instructional minutes, F (5,96) = 1.26, $p > .05$. Because the analysis of variance for instructional time showed no significant differerces, further pairwise comparisons of group means for this variable were abandoned.

Five non-orthogonal pairwise comparisons were of interest. Multiple t ratios were used for calculating these planned comparisons:

C1: The group receiving "increasing ratio review" (group 1) made significantly fewer DEs on the posttest than did the group receiving "immediate review only" (group 2), (t_{obt} = 3.214) > (t_{crit} = 1.985, df = 96, alpha = .025).

Similarly, the increasing ratio review group made significantly fewer OLEs on the posttest than did the immediate review only group, (t_{obt} = 3.818) > (t_{crit} = 1.985, df = 96, alpha = .025).

C2: The group receiving "immediate review only" (group 4) did not make significantly fewer DEs on the posttest than did the group receiving "no review" (group 5), (t_{obt} = 0.109) < (t_{crit} = 1.985, df = 96, alpha = .025).

Similarly, the immediate review only group did not make significantly fewer OLEs on the posttest than did the no review group, (t_{obt} = 0.252) < (t_{crit} = 1.985, df = 96, alpha = .025).

C3: The group receiving "adaptive feedback" (group 3) did not make significantly fewer DEs on the posttest than did the group receiving "fixed feedback" (group 6), (t_{obt} = 0.018) < (t_{crit} = 1.985, df = 96, alpha = .025).

Similarly, the adaptive feedback group did not make significantly fewer OLEs on the posttest than did the fixed feedback group, (t_{obt} = 0.191) < (t_{crit} = 1.985, df = 96, alpha = .025).

C4: In the presence of increasing ratio review, the group receiving "discrimination training" (group 1) made significantly fewer DEs on the posttest than did the group receiving "no discrimination training" (group 3), (t_{obt} = 2.251) > (t_{crit} = 1.985, df = 96, alpha = .025).

However, there were no significant differences between the two groups in the mean number of OLEs on the posttest, (t_{obt} = 0.090) < (t_{crit} = 1.985, df = 96, alpha = .025).

C5: The group receiving "adaptive feedback and discrimination training" (group 1) made significantly fewer DEs on the posttest than did the group receiving "fixed feedback and no discrimination training" (group 6), (t_{obt} = 2.269) > (t_{crit} = 1.985, df = 96, alpha = .025).

However, there were no significant differences between the two groups in the mean number of OLEs on the posttest, (t_{obt} = 0.281) < (t_{crit} = 1.985, df = 96, alpha = .025).

An analysis of frequency and type of errors during instruction showed that there were no significant differences in these measures between groups. Nearly 42% of the errors during instruction were DEs. Analysis of covariance (with the S's pretest and vocabulary test scores as covariates) showed results identical to the above for all comparisons.

Discussion

These results suggest that for computerized drills:

1. Increasing ratio review techniques (e.g., 1-, 3-, 5-later review) are more effective than traditional review techniques (e.g., 1-later review) or no review at all. Moreover, 1-later review is

no more effective than no review. This suggests that a 2-, 3-, 5-later review schedule may have been more effective than a 1-, 3-, 5-later review schedule.

2. Although adaptive feedback without discrimination training was no more effective than fixed feedback in reducing the number of discrimination errors on the posttest (group 3 vs. group 6), adaptive feedback with discrimination training significantly reduced the number of discrimination errors on the posttest (group 1 vs. group 3; group 1 vs. group 6). Thus, the effectiveness of adaptive feedback may be significantly lessened unless it is accompanied by discrimination training. For example, suppose a S confused "aunt-obasan" with "mother-okaasan." The student typed "okaasan" for "aunt." If the S was in the "adaptive" feedback group, the S received the DE feedback message: "okaasan is mother; aunt is obasan." If the S was never required to discriminate between the stimuli "aunt" and "mother" when presented simultaneously, however, (i.e., if there was no discrimination training), the S may not have been reinforced for attending to the first part of the DE feedback message ("okaasan is mother"). Thus, the S's attention may have been "short-circuited" since the S had to remember only the second part of the DE feedback message ("aunt is obasan") to successfully translate a future review presentation of "aunt;" this short-circuited message was identical to the fixed feedback message.

Note that among these three groups (groups 1, 3, and 6) there were no significant differences in the mean number of OLEs on the posttest. Of course this is what you would expect because adaptive feedback and fixed feedback treat OLEs in the same way.

Frequently in non-drill computer lessons the student is required to type a missed answer immediately following corrective feedback. This requirement may be less than useful in facilitating learning.

The author suspects that the student learns to type the correct answer without attending to the question paired with the answer.

3. Finally, the additional presentations inherent in employing increasing ratio review and discrimination training techniques did not significantly add to the instructional time. For example, the mean number of item presentations for group 1 (adaptive feedback, discrimination training, and increasing ratio review) was 288 items (SD = 129); the mean number of item presentations for group 5 (adaptive feedback, no discrimination training, no review) was 158 items (SD = 72). Yet, there were no significant differences between these two groups in instructional time. Thus, group 1 completed more items per minute during instruction than did group 5. An explanation for this result may be found in the group means for "time on posttest." An analysis of covariance (with "time on the typing test" as the covariate) showed that group 1 (adaptive feedback, discrimination training, and increasing ratio review) completed the posttest significantly faster than any other group, suggesting that the Ss in group 1 were more "confident" of their answers. This confidence hypothesis also may explain why the instructional time for group 1 was not significantly different from any other group. This is an important finding. More educationally effective procedures can be implemented without spending additional time in instruction.

What conclusions can be made from these findings that would assist instructional designers in implementing more effective computerized drills? Most important, increasing ratio review techniques should be employed. Exactly what specific review schedules are most beneficial and under what instructional settings remains a question for future research. Second, if a drill list includes stimulus-response pairs that may be confused with other stimulus-response pairs in the list, the drill should include adaptive feedback procedures. Discrimination training should be provided after discrimination error feedback so that the student has an opportunity to "work out" the confusion.

Instructional Demonstrations of the CFP

The CFP was designed originally in 1974. Since then it has been used in a number of diverse instructional settings on the PLATO system. Three of these applications will be described below, illustrating the CFP's flexibility. Additional considerations in the design of computerized drill lessons will be explained as well.

U.S. and Illinois Constitution Drill

A requirement for obtaining a high school general equivalency diploma (GED) in the State of Illinois is the passing of a multiple choice exam on the U.S. and Illinois constitutions, U.S. and Illinois history, and the U.S. flag code. The exam is administered by the State of Illinois, and students often would pass the GED exam, yet fail the "constitution test." This was a particular problem in Illinois prisons where PLATO was being used for GED instruction (Siegel, 1978). We, therefore, were encouraged by the instructors to develop a PLATO drill and practice program for their students.

We began with the extensive worksheets that the instructors had already prepared for classroom instruction. These consisted of multiple-choice and matching tasks. Because we were interested in comparing classroom instruction with the CFP drill instruction, we used the same materials for the development of the drill lesson.

The lesson (Siegel, Felty, and Reichmann, 1979) was structured as follows:

There were six drills and seven tests. The first six tests covered the material in the drills; the seventh test was a composite mastery test consisting of the 60 most difficult items from the previous six tests (based on student performance data at the time of the testing). The

student began instruction by taking the first drill. Twenty-four hours after completing the drill, the student was allowed to take the corresponding test. If the student scored less than 85% on the test, the student was required to complete the drill again. The corresponding test was not available again until 24 hours after completion of the drill. In this way students could determine their long term retention of the material presented in each drill. Upon reaching criterion for each drill and test, the student was allowed to take the composite mastery test.

The first drill covered U.S. constitution vocabulary. At the top of the screen the following question appeared: "Which phrase below does this statement describe?" An example statement was "a counting of the people living in the United States." Below this statement appeared a list of terms arranged in alphabetical order:

h ammendment
c apportionment
f bicameral
m bill of attainder
j census
o compromise
i concurrent
q delegated powers
l eminent domain
a ex post facto
s executive department
p federal system of government
e habeas corpus
r impeachment
k implied powers
b judicial department
n legislative department
d president pro tempore
g reserved powers

The student pressed the adjacent letter to indicate his response. Note that the letters were not presented in alphabetical order. In fact, after each item the adjacent letters were randomly rearranged. This "forced" the student to attend to the vocabulary words rather than to the corresponding

letters. If the letters were always fixed, the student might have learned the interesting but useless paired associate: "a counting of the people living in the United States = j."

The drill employed the usual CFP features: (a) increasing ratio review (2-, 4-, 6-later review); (b) adaptive feedback; and (c) discrimination training. When the student chose a response from the list of terms, an arrow appeared adjacent to the selected term. If the selected term was correct, praise was provided (e.g., "ok"); if the selected term was incorrect, a box flashed around the correct term. In this drill all errors are DEs rather than OLEs by definition. The other vocabulary and multiple-choice drills were handled in a similar fashion.

Although the prison environment did not allow us to run a carefully controlled experiment, we were able to evaluate the drill procedures. Students were equally divided between the PLATO group and the classroom only group. The students were not randomly assigned to the treatment groups, but typically the higher performing students were assigned to the classroom treatment.

The passing score on the test was 30. The following results were recorded for the 77 Classroom and 80 PLATO CFP students who took the test. The average scores on the test were:

Classroom	29 points
PLATO CFP	40 points

while the average amounts of time spent in learning were:

Classroom	16 hours
PLATO CFP	4 hours

Although this was an informal field study rather than a carefully controlled laboratory experiment, these differences were so large that they

were almost certainly real. In fact, the instructors and students who used the lessons regarded their superiority as obvious.

The Decode Drill

Usually computerized drills involve a student and a computer. The "Decode Drill" (Siegel & Fuller, 1978) involves a human tutor as well. The student and tutor sit in front of the computer screen. Single words are displayed on the screen for the student to decode or sound-out. The tutor provides the computer's "ears" in this beginning reading drill. The words are stored in a database, and the complete list of words is subdivided into smaller sublists -- one sublist for each drill session. If the student correctly reads the word, the tutor praises the student and presses the "NEXT" key on the keyset. If the student incorrectly reads the word, the tutor corrects the student by saying, for example, "House. What word?" The student reads the word again and the tutor presses the "SHIFT-NEXT" key. This signals to the drill lesson that the student made a mistake. Then the queue of words in the drill session is automatically rearranged according to an increasing ratio review schedule. (The tutor's directions and script unobtrusively appears on the screen as well.) The drill session ends when the student has successfully decoded each word three times without an intervening mistake (excluding correct responses on the first two review positions). Succeeding drill sessions begin with a new set of words as well as a few of the most difficult words (for that student) from the previous drill session.

Although not available in the current implementation of the decode drill, it would be possible to employ discrimination feedback and training. For example, the student reads "horse" for "house." Immediately the tutor would type the word "horse." As it appears on the screen below the word "house," the tutor would say, "You read this word (a box would flash around 'horse')... but I want you to read this word (the box would then flash around 'house')." If the student continues to make a mistake, the tutor

would read the word "house." But now the word "horse" has been entered into the database and discrimination training could commence with missed and confused words.

The advantage of these drill procedures is that the hearing and speech capabilities of a human tutor are combined with the sophisticated drilling techniques of a computer in a way which neither alone could successfully implement. Moreover, the tutor need not be a classroom teacher; the tutor could be another student who has mastered the list of words. It even may be possible to replace the human tutor with a synthetic speech device that would provide auditory feedback to the student. If the student judged that he correctly decoded the word, he would press the NEXT key; however, if he judged that he incorrectly decoded the word, he would press the SHIFT-NEXT key. Again, the computer would control the sequencing of items.

The Reading Comprehension Drill

Typically computerized drills have been used for teaching math facts or foreign language vocabulary. But the PCP Reading Comprehension Program (Siegel, DiBello, Bruner, & Felty, 1977) uses a computerized drill to teach the relationship between WH-questions (who, what, when, where, why, how, etc.) and sentences which answer these questions. Within the scope of this paper it is not feasible to describe each of the 25 teaching formats which have been developed (see Siegel, 1979). Rather, one of these formats is reproduced below.

The student is presented with a sentence or a passage and is asked to underline the part that tells the answer to a given question. For example:

Read this:

> The first knitting machine was invented in 1589 in England, and the first cotton gin was invented in 1793 in America.

Then the student is asked to draw a line under the part that tells "where the first cotton gin was invented." The student presses the "NEXT" key on his keyset, and a line is automatically drawn under consecutive phrases of the sentence -- one phrase at a time. In this sentence, the first time the student presses NEXT, the phrase "The first" is underlined; the second time the student presses NEXT, the previous line is erased and the phrase "knitting machine" is underlined. In this manner the student cycles through the different phrases in the sentence. (Notice that the student does not have to worry about the length of the phrase or what words are included in the phrase; the format automatically handles this.) When the student has underlined his answer, he presses the "LAB" key.

If a student makes a mistake on this item, the CFP procedures are used. For example, a DE initiates a message such as: "No. That part tells <u>when</u> the first cotton gin was invented. Underline the part that tells <u>where</u> the first cotton gin was invented." This discrimination problem -- confusing "when" and "where" questions -- is practiced on succeeding tasks in the drill; that is, additional tasks are automatically inserted into the lesson which work on the when/where confusion. If the student cannot answer the original question after being given DE feedback, the student is told the correct answer (OLE feedback):

Look:

The first knitting machine was invented in 1589 in England, and the first cotton gin was invented in 1793 [where] ?

The first knitting machine was invented in 1589 in England, and the first cotton gin was invented in 1793 <u>in America</u> .

An interesting feature of this format is that the item will be slightly changed when it reappears in the lesson for review; this discourages

memorization of the correct answer. In the above item, for example, there are four phrases that may change: (a) "in 1589" may be replaced by "in the sixteenth century," "in the fifteen hundreds," "almost 400 years ago," or "in the 1580's;" (b) "in England" may be replaced by "in Europe," "in the British Isles," or "in Great Britain;" (c) "in 1793" may be replaced by "soon after the Revolutionary War," "in the late 1700's," "in the eighteenth century," or "shortly before 1800;" and (d) "in America" may be replaced by "in the United States," "in our country," or "on this side of the Atlantic." These semantically equivalent phrases or "flip-book" parts greatly increase the number of different items that the student can read. This feature, which could only be implemented on a computer, allows for the generation of more than 200 million different sentences throughout the twenty reading comprehension lessons!

Unlike static paired associate items (e.g., "Si / silicon," "la fleur / flower," "3 + 2 = ? / 5"), the above reading item changes from presentation to presentation in a non-random way. Namely, if the student makes a DE -- e.g., confusing "when the knitting machine was invented" with "when the cotton gin was invented" -- the drill will systematically vary these flip-book parts on succeeding review tasks (e.g., replacing "in 1589" with "in the sixteenth century"):

Original sentence:

> The first knitting machine was invented in 1589 in England, and the first cotton gin was invented in 1793 in America.

Varying the first "when":

> The first knitting machine was invented in the 16th century in England, and the first cotton gin was invented in 1793 in America.

Varying the second "when":

> The first knitting machine was invented in 1589 in England, and the first cotton gin was invented shortly before 1800 in America.

Varying the first and second "when":

> The first knitting machine was invented in the late 1500's in England, and the first cotton gin was invented in the 18th centur in America.

Varying the sentence structure:

> Almost 400 years ago the first knitting machine was invented in England, and shortly before 1800 the first cotton gin was invented in America.

The student must focus on the structure of the sentence rather than on a particular memorized phrase. Extending this idea to the teaching of concepts, a drill item can represent an individual concept and a drill session can represent a collection of concepts (or a subset of concepts in a conceptual hierarchy). For each item in the drill (an instance of the concept), critical features can be held constant and irrelevant features can vary from instance to instance (successive presentations of the concept in the drill). DEs can initiate procedures which systematically vary a particular instance of the concept, thus training the student to attend to the concept's relevant features.

Designs for the Future

The mathematical learning model and optimal control strategies employed by Atkinson (1972b) as well as the instructional strategies employed by Siegel and Misselt (1974) should be viewed as first steps in the design of optimal drill paradigms. With the great expansion in computer availability it is important to develop a base of theory together with a set of applications guidelines that will greatly ease the task of future instructional designers and programmers in the development of drill and practice instruction. Hopefully much of the trial-and-error work that has been required to set up the CFP in a given instructional drill situation will be obviated. (See Siegel and DiBello, 1979.)

The use of drills in areas other than simple paired-associate tasks needs to be fully explored. As drill paradigms are more widely viewed as

potential instructional tools for the teaching of concepts and sophisticated cognitive networks, drill paradigms will become central to the design of instructional lessons, and not merely an appended technique that is used to "firm-up" what has already been taught.

Footnotes

[1] In actual practice, there are some additional corrective feedback considerations: (a) How will spelling errors be treated? In particular, how will misspelled OLEs (e.g., "mettal") or misspelled DEs (e.g., "silvir") be corrected? For purposes of this paper we will treat all spelling errors as OLEs, but there are other options available to the designer. (b) Should all OLEs be corrected in the same way? For example, one could imagine that an appropriate feedback message to the student who responded "metal" to the stimulus "Si" might be: "Si is a metal. But what is the name for this metal?" This is a form of DE feedback. In fact, any incorrect response (excluding nonsense words) could be considered as a DE. But for simplicity, we will limit DEs to responses which are answers to other items in the drill list. One would not program special feedback messages unless the student response data indicated that there was a special discrimination problem that needed attention.

References

Atkinson, R. C. Ingredients for a theory of instruction. American Psychologist, 1972a, 27, 921-931.

Atkinson, R. C. Optimizing the learning of a second-language vocabulary. Journal of Experimental Psychology, 1972b, 96, 124-129.

Atkinson, R. C., & Crothers, E. J. A comparison of paired-associate learning models having different acquisition and retention axioms. Journal of Mathematical Psychology, 1964, 1, 285-315.

Boltyanskii, V. G. [Mathematical methods of optimal control.] (I. Tarnove, Ed. & K. Trirogoff, Trans.). New York: Holt Rinehart & Winston, Inc., 1971.

Calfee, R. C., & Atkinson, R. C. Paired-associate models and the effects of list length. Journal of Mathematical Psychology, 1965, 2, 254-265.

Greeno, J. G. Paired-associate learning with massed and distributed repetitions of items. Journal of Experimental Psychology, 1964, 67, 286-295.

Hintzman, D. L. Theoretical implications of the spacing effect. In R. L. Solso (Ed.), Theories in Cognitive Psychology: The Loyola Symposium. New York: Wiley, 1974.

Kalman, R. E., Falb, P. L., Arbib, M. A. Topics in mathematical systems theory. New York: McGraw Hill Book Co., 1969.

Keppel, G. Facilitation in short- and long-term retention of paired associates following distributed practice in learning. Journal

of Verbal Learning and Verbal Behavior, 1964, 3, 91-111.

Kulhavy, R. C. Feedback in written instruction. Review of Educational Research, 1977, 47(2), 211-232.

Peterson, L. R. Short-term verbal memory and learning. Psychological Review, 1966, 73, 193-207.

Peterson, L. R., Wampler, R., Kirkpatrick, M., & Saltzman, D. Effect of spacing presentations on retention of a paired associate over short intervals. Journal of Experimental Psychology, 1963, 66, 206-209.

Siegel, M. A. Computer-based education in prison schools. Journal of Educational Technology Systems, 1978-79, 7(3), 239-256.

Siegel, M. A. Teacher's guide to the PCP reading comprehension program for adolescent and adult learners (CERL Report E-3). Urbana, Il.: Computer-based Education Research Laboratory, January 1979.

Siegel, M. A., & DiBello, L. V. Optimizing the performance of computerized drills. Proposal submitted to the Army Research Institute. Urbana, Il.: Computer-based Education Research Laboratory, January 1979.

Siegel, M. A., DiBello, L. V., Bruner, E. C., Felty, J. M. PCP reading comprehension I-IV. Urbana, Il.: Computer-based Education Research Laboratory, 1977.
(PLATO lesson)

Siegel, M. A., Felty, J. M., Reichmann, L. P. United States and Illinois constitutions, history and government, declaration of independence, and United States flag code: Drills and tests. Urbana, Il.: Computer-based Education Research Laboratory, 1979.
(PLATO lesson)

Siegel, M. A., & Fuller, D. Decode. Urbana, Il.: Computer-based Education Research Laboratory, 1978. (PLATO lesson)

Siegel, M. A., & Misselt, A. L. A corrective feedback and criterion teaching paradigm for CAI. Unpublished manuscript, University of Illinois, 1974.

Smith, S. G., & Sherwood, B. A. Educational uses of the PLATO computer system. Science, 1976, 192, 344-352.

Young, J. L. Reinforcement-test intervals in paired-associate learning. Journal of Mathematical Psychology, 1971, 8, 58-81.